Parameter Identification for a Dispersive Dielectric in 2D Electromagnetics: Forward and Inverse Methodology with Statistical Considerations.

H. T. Banks†

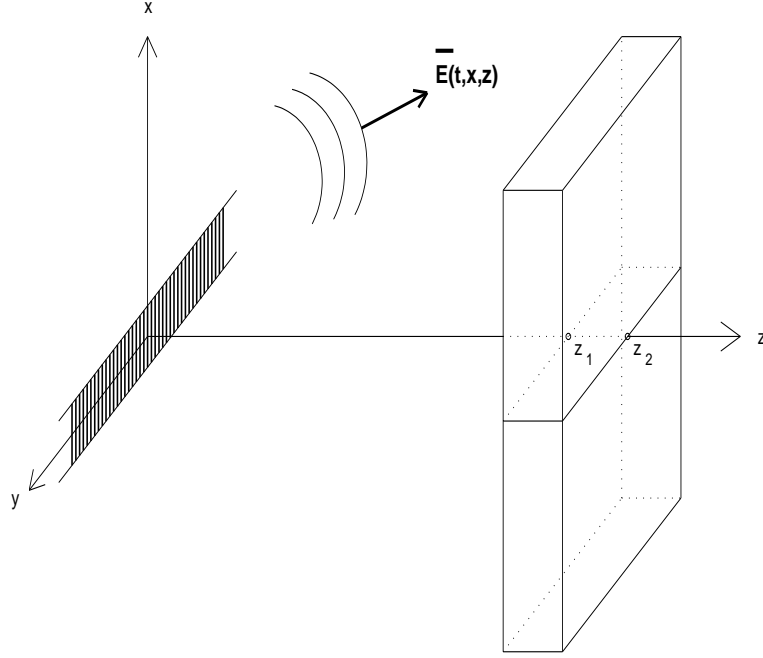† Center for Research in Scientific Computation, North Carolina State University, Raleigh, NC 27695-8205

J. M. Bardsley‡

‡ Department of Mathematical Sciences, University of Montana, Missoula, MT 59812-0864

16 December 2003

Abstract. We present methodology for obtaining forward solutions to Maxwell's equations in two dimensions, in the presence of a Debye medium. Perfectly Matched Layer (PML) absorbing boundary conditions are used to absorb incoming energy at the finite boundaries. A time-domain, PDE formulation is presented, and a finite difference time-domain (FDTD) algorithm is used to obtain numerical solutions. A least squares formulation of the inverse problem results from a careful consideration of the noise model for data generation. The inverse problem is solved with varying levels of noise in the data, and a frequency domain analysis is given that provides an explanation of the results. The results and analysis motivate strategies for solving the inverse problem that decrease computational cost. Finally, a result from the statistical theory of large samples is used to obtain estimates of the variability in parameter estimates that is due to the variability in the noisy data.

Report Documentation Page

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
16 DEC 2003		**00-00-2003 to 00-00-2003**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Parameter Identification for a Dispersive Dielectric in 2D Electromagnetics: Forward and Inverse Methodology with Statistical Considerations	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
North Carolina State University,Center for Research in Scientific Computation,Raleigh,NC,27695-8205	

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
see report

15. SUBJECT TERMS

| 16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**25**	

Figure 1. Infinite strip antenna and dielectric slab.

1. Introduction

In [3] the authors developed techniques for solving Maxwell's equations in one dimension in a Debye medium, and for solving the associated parameter estimation or inverse problem. The experimental setup assumed in the formulation of the one dimensional problem includes an antenna that has infinite extension in both the x and y directions (i.e., an infinite sheet). In this paper, we consider instead the case of an infinite strip antenna which is infinite in the y direction, but is finite in the x direction. This setup is illustrated schematically in Figure 1. The resulting Maxwell system involves oblique incident waves on the dielectric slab and hence is two dimensional. The goal of this paper is to extend results in [3] to this problem.

Regardless of the experimental setup, in practice absorbing boundary conditions are necessary in order to model the (effectively) infinite spatial domain by a finite computational domain. In the one dimensional case, perfectly absorbing boundary conditions exist and are easily implemented [3]. Unfortunately, such boundary conditions do not exist in higher dimensions. Though various analytic "absorbing" boundary conditions are available for use on the two dimensional problem, a more effective technique employs perfectly matched layers (PMLs) as a fictitious absorbing layer surrounding the region of interest (see Figure 2). If the PML is constructed properly, there is no reflection at the PML interface and any energy returning to the domain of interest after travelling through the PML is negligible. In Section 2, we present a system of PDEs that models the setup illustrated in Figure 1 with a Debye dielectric slab and PML absorbing boundaries. This system assumes that the electric
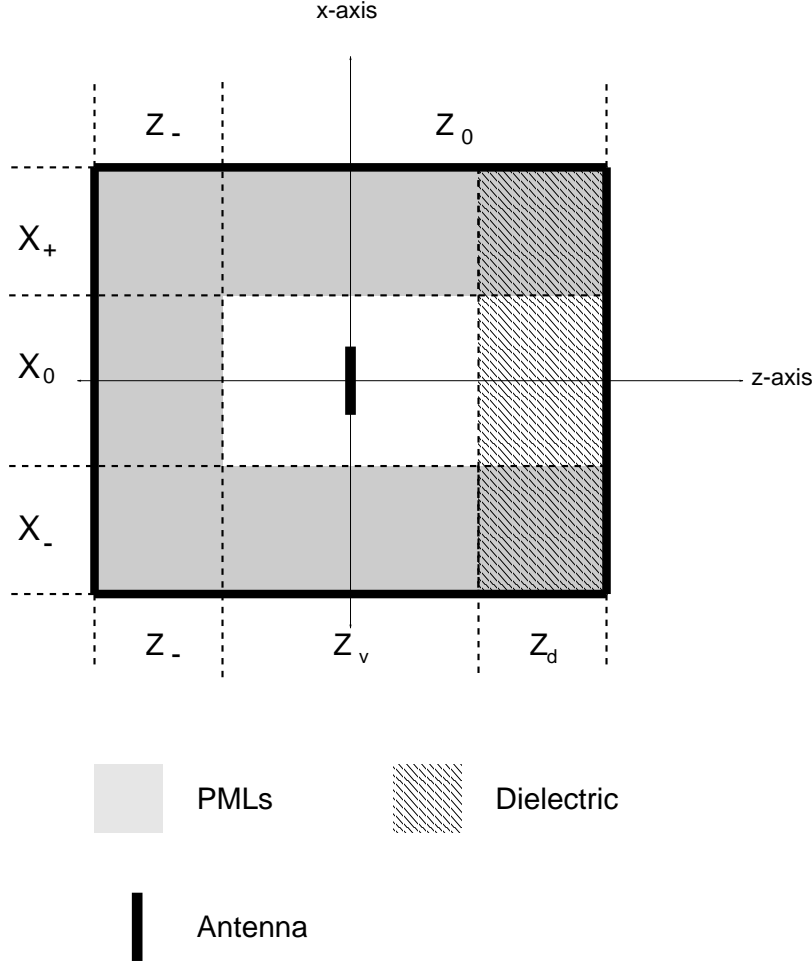
Figure 2. Problem geometry

field generated by the antenna is polarized so that it only has a y component and is solved using a finite difference time-domain (FDTD) algorithm [16]. Numerical results from forward simulations are presented.

In Section 3 we discuss the solution of the associated parameter identification problem. That is, given data generated from the experimental setup illustrated in Figure 1 with a Debye dielectric slab, determine the parameters that characterize the Debye medium. We begin the section with the noise model for data generation for such experiments. Using this noise model, we are led to the negative log-likelihood function, which takes a least squares form. The inverse problem is then to minimize this function. We discuss the modified Levenburg-Marquardt method that we used to this end. Several attempts are made at the inverse problem with varying levels of noise. The results are explained and new approaches are motivated by a frequency domain analysis. This analysis allows one to obtain a knowledge of which parameters are or are not likely to be reconstructed accurately.

Because our data is noisy, it can be viewed as a realization of a random variable and is therefore variable. That is, in the context of the setup of Figure 1, if we perform exactly the same experiment many times, we will obtain different data sets each time. Consequently, if we solve the inverse problem with each of these different data sets, we will obtain different values for the reconstructed parameters. Thus the variability in the data translates into variability in the parameter estimates. In Section 4, we use the statistical theory of large samples [9] to obtain estimates for the variability in our estimates. Conclusions are presented in Section 5

2. The Forward Problem

We begin with Maxwell's equations for a region with free charge density $\rho = 0$, which are given by

$$\nabla \cdot \mathbf{D} = 0 \tag{1}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2}$$

$$\nabla \times \mathbf{E} = -\partial_t \mathbf{B} \tag{3}$$

$$\nabla \times \mathbf{H} = \partial_t \mathbf{D} + \mathbf{J}, \tag{4}$$

where the vectors in (1)-(4) are functions of position $\mathbf{r} = (x, y, z)$ and time t, and $\mathbf{J} = \mathbf{J}_c + \mathbf{J}_s$, where $\mathbf{J}_c$ is the conduction current density and $\mathbf{J}_s$ is the source current density. We assume only free space can have a source current, and thus either $\mathbf{J}_c = \mathbf{0}$ or $\mathbf{J}_s = \mathbf{0}$, depending on whether or not the region is free space. For a Debye medium, magnetic effects are neglected, and we assume that Ohm's law governs the electric conductivity. Hence, within the dielectric medium

$$\mathbf{B} = \mu_0 \mathbf{H}, \tag{5}$$

$$\mathbf{J}_c = \sigma \mathbf{E}. \tag{6}$$

The displacement vector $\mathbf{D}$, on the other hand, has a more complex representation, namely,

$$\mathbf{D} = \epsilon_0 \epsilon_\infty \mathbf{E} + \mathbf{P}, \tag{7}$$

where $\mathbf{P}$ is the electric polarization vector that satisfies the differential equation

$$\tau \dot{\mathbf{P}} + \mathbf{P} = \epsilon_0 (\epsilon_s - \epsilon_\infty) \mathbf{E} \tag{8}$$

inside of the dielectric and is $\mathbf{0}$ outside of the dielectric.

The computational domain for our problem is given in Figure 2. Notice first that it is contained in the x,z plane. This is facilitated by the fact that we have assumed uniformity along the y-axis. Secondly, note that the computational domain has finite support. This is desirable since in order to numerically solve Maxwell's equations, a finite computational grid is necessary. For our problem, the resulting computational boundaries will generally reflect electromagnetic waves, and hence, numerical solutions will contain non-physical artifacts. One of the more effective ways to combat this difficulty is to use Perfectly Matched Layers (PMLs). The PML is a fictitious medium

that is placed on the outside of the region of interest (see Figure 2) in order to absorb incoming energy without reflection.

In order to be more explicit about the geometry of the computational domain, we introduce some notation. Let $\mathcal{D}$ denote the computational domain as pictured in Figure 2. We partition the interval X into disjoint closed intervals X_-, X_0, and X_+ such that

$$\max X_- = \min X_0$$
$$\max X_0 = \min X_+,$$

and partition the interval Z into disjoint closed intervals Z_-, Z_v, and Z_d such that

$$\max Z_- = \min Z_v$$
$$\max Z_v = \min Z_d.$$

The regions Z_v and Z_d are the vacuum and dielectric regions respectively. Let

$$Z_0 = Z_v \cup Z_d.$$

The region $X_0 \times Z_0$ will be our domain of interest. The buffer region $\mathcal{D} \setminus (X_0 \times Z_0)$ outside our domain of interest contains the PMLs. Note that the PMLs surround the computational region on only three sides. This suffices since it is assumed that the dielectric is backed by a reflective material such as metal, where the boundary condition $\mathbf{E} = 0$ is used.

Our source term $\mathbf{J}_s$ models a finite antenna in free space. It will have the form

$$\mathbf{J}_s(t, x, z) = \mathcal{I}_{[-\bar{x}, \bar{x}]}(x) \cdot \delta_{\bar{z}}(z) \cdot \begin{bmatrix} f_x(t) \\ f_y(t) \\ f_z(t) \end{bmatrix} \tag{9}$$

where $\mathcal{I}_{[a,b]}$ is the indicator function on the interval $[a, b]$, $[-\bar{x}, \bar{x}] \in X_0$, $\bar{z}$ is the z-coordinate of the position of the antenna. The functions f_m, for $m = x, y, z$, are determined by the polarization and intensity profile of the interrogating electromagnetic pulse.

We now express Maxwell's equations in a form that is useful for our particular problem. Combining the curl equations (3) and (4) with equations (5), (6), (7), and (8), using the rescaling

$$\widetilde{\mathbf{E}} = \sqrt{\frac{\epsilon_0}{\mu_0}} \mathbf{E}, \tag{10}$$

and writing the resulting equations in the frequency domain, we have

$$j\omega \widehat{\widetilde{\mathbf{D}}} = c_0 \cdot \nabla \times \widehat{\mathbf{H}} - \widehat{\mathbf{J}}_s \tag{11}$$

$$\widehat{\widetilde{\mathbf{D}}}(\omega) = \epsilon_r^*(\omega) \cdot \widehat{\widetilde{\mathbf{E}}}(\omega) \tag{12}$$

$$j\omega \widehat{\mathbf{H}} = -c_0 \cdot \nabla \times \widehat{\widetilde{\mathbf{E}}}, \tag{13}$$

where "^" denotes the Fourier transform in the time variable, and $\epsilon_r^*(\omega)$ is given by

$$\epsilon_r^*(\omega) = \epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 + j\omega\tau} + \frac{\sigma}{j\omega\epsilon_0}. \tag{14}$$

Note that the conductivity term $\mathbf{J}_c = \sigma\mathbf{E}$ is now contained within (12) rather than within the forcing term in (11), as was the case in the original formulation (1)-(4). In particular, it appears in the last term of (14). The polarization equation (8) is also contained with (12). It is found in the second term of (14). We express Maxwell's equations in this way for three reasons. First, in this formulation, the parameters that we hope to identify in the inverse problem are contained entirely within the single equation (12) with (14). Secondly, given particular values for the parameters ϵ_∞, ϵ_s, τ, and σ, for a fixed frequency ω, equation (14) will reveal the sensitivity of the displacement vector $\widehat{\widetilde{\mathbf{D}}}$ to changes in these parameters. This information will be particularly useful when we attempt the parameter identification problem. Third, this formulation results in a simplified analysis, PML formulation, and PML implementation.

2.1. Reduction to Two Dimensions and to the TM Mode

Since the problem geometry does not depend on y, Maxwell's equations can be simplified. In addition, we make the assumption that our pulse is polarized so that it only has a y component. We define f_y in (9) by

$$f_y(t) = \mathcal{I}_{[0,t_f]}(t)sin(\hat{\omega}t), \tag{15}$$

where $\hat{\omega}t_f$ is an integral multiple of 2π, and hence,

$$\mathbf{J}_s(t,x,z) = \begin{bmatrix} 0 \\ \mathcal{I}_{[-\bar{x},\bar{x}]}(x) \cdot \delta_{\bar{z}}(z) \cdot \mathcal{I}_{[0,t_f]}(t) \cdot sin(\hat{\omega}t) \\ 0 \end{bmatrix}. \tag{16}$$

The curl equations (11),(13) then have the equivalent time-domain formulation

$$\begin{bmatrix} \partial_t H_x(x,z) \\ \partial_t H_y(x,z) \\ \partial_t H_z(x,z) \end{bmatrix} = c_0 \begin{bmatrix} \partial_z \widetilde{E}_y(x,z) \\ \partial_x \widetilde{E}_z(x,z) - \partial_z \widetilde{E}_x(x,z) \\ -\partial_x \widetilde{E}_y(x,z) \end{bmatrix}, \tag{17}$$

$$\begin{bmatrix} \partial_t \widetilde{D}_x(x,z) \\ \partial_t \widetilde{D}_y(x,z) \\ \partial_t \widetilde{D}_z(x,z) \end{bmatrix} = c_0 \begin{bmatrix} -\partial_z H_y(x,z) \\ \partial_z H_x(x,z) - \partial_x H_z(x,z) \\ \partial_x H_y(x,z) \end{bmatrix} - \mathbf{J}_s(t,x,z), \tag{18}$$

where $\mathbf{J}_s$ is given by (16).

Assuming zero initial conditions for $\mathbf{E}$ and $\mathbf{H}$, the vector system (17)-(18) reduces to the set of three equations

$$\frac{\partial \widetilde{D}_y}{\partial t} = c_0 \left(\frac{\partial H_z}{\partial x} - \frac{\partial H_x}{\partial z} \right) - J_{s,y}(t,x,\bar{z}), \tag{19}$$

$$\frac{\partial H_x}{\partial t} = c_0 \frac{\partial \widetilde{E}_y}{\partial z}, \tag{20}$$

$$\frac{\partial H_z}{\partial t} = -c_0 \frac{\partial \widetilde{E}_y}{\partial x}, \tag{21}$$

which is known as the TM mode.

We choose to work with the TM mode since (11)-(13) then reduces to four equations. If we chose instead to polarize our electromagnetic pulse so that $J_{s,y} = 0$, while $J_{s,x}, J_{s,z} \neq 0$, the TE mode would result, and (11)-(12) would reduce instead to five equations. In the TM case, (12) is needed only for updating E_y, whereas in the TE case it is used to update both E_x and E_z.

2.2. Perfectly Matched Layers

To simplify notation in the sequel, we suppress the "~" notation found in (19)-(21). In order to give physical motivation for the use and effectiveness of the PML we first consider (19). We note that $J_{s,y} = 0$ inside of the PML regions. To model a lossy medium for $z \in Z_-$ (see Figure 2), we add a loss term:

$$\frac{\partial D_y}{\partial t} = c_0 \left(\frac{\partial H_z}{\partial x} - \frac{\partial H_x}{\partial z} \right) - \frac{\hat{\sigma}(z)}{\epsilon_0} D_y, \tag{22}$$

where $\hat{\sigma}(z) = 0$ for $z \in Z_0$. Inside of the PML, i.e. for $z \in Z_-$, $\hat{\sigma}(z)$ is a smooth function of z that increases from a value of zero at the free-space/PML interface to a value of σ_{max} at the boundary. In the frequency domain, equation (22) can be written as

$$j\omega \cdot \left(1 + \frac{\hat{\sigma}(z)}{j\omega\epsilon_0} \right) \widehat{D}_y = c_0 \left(\frac{\partial \widehat{H}_z}{\partial x} - \frac{\partial \widehat{H}_x}{\partial z} \right). \tag{23}$$

Making similar choices for (20) and (21), we obtain

$$j\omega \cdot \left(1 + \frac{\hat{\sigma}(z)}{j\omega\epsilon_0} \right) \widehat{D}_y = c_0 \left(\frac{\partial \widehat{H}_z}{\partial x} - \frac{\partial \widehat{H}_x}{\partial z} \right), \tag{24}$$

$$j\omega \cdot \left(1 + \frac{\hat{\sigma}(z)}{j\omega\epsilon_0} \right)^{-1} \widehat{H}_x = c_0 \frac{\partial \widehat{E}_y}{\partial z}, \tag{25}$$

$$j\omega \cdot \left(1 + \frac{\hat{\sigma}(z)}{j\omega\epsilon_0} \right) \widehat{H}_z = -c_0 \frac{\partial \widehat{E}_y}{\partial x}, \tag{26}$$

where the "^" over E_y, D_y, H_x, and H_z denotes the Fourier transform in time variable.

For $z \in Z_-$, the PML is modelled in the frequency domain by (24)-(26). These very specific choices for the fictitious complex permittivity $\hat{\epsilon}_y(z) = 1 + \hat{\sigma}(z)/j\omega\epsilon_0$ and permeabilities $\hat{\mu}_x(z) = (1 + \hat{\sigma}(z)/j\omega\epsilon_0)^{-1}$, $\hat{\mu}_z(z) = 1 + \hat{\sigma}(z)/j\omega\epsilon_0$ result in zero reflection (in the continuum space) of electromagnetic energy at the free-space/PML interface [15]. To see this, we note that if a wave is propagating in a medium A and impinges upon a medium B, the amount of reflection is determined by the respective impedances and is given by

$$\Gamma = \frac{\eta_A - \eta_B}{\eta_A + \eta_B}. \tag{27}$$

The impedance $\hat{\eta}$ of a medium with complex permittivity $\hat{\epsilon}$ and complex permeability $\hat{\mu}$ is given by

$$\hat{\eta} = \sqrt{\frac{\hat{\mu}}{\hat{\epsilon}}}.$$

For our formulation, $\hat{\epsilon}$, $\hat{\mu}$, and $\hat{\eta}$ are diagonal tensors of the form

$$\hat{\epsilon} = \text{diag}(\hat{\epsilon}_x, \hat{\epsilon}_y, \hat{\epsilon}_z),$$
$$\hat{\mu} = \text{diag}(\hat{\mu}_x, \hat{\mu}_y, \hat{\mu}_z),$$
$$\hat{\eta} = \text{diag}(\hat{\eta}_x, \hat{\eta}_y, \hat{\eta}_z).$$

The dielectric constants $\hat{\epsilon}_y$, $\hat{\mu}_x$, and $\hat{\mu}_z$ are chosen so that they satisfy two sets of conditions, one of which is that the resulting diagonal tensor Γ defined by (27) is the zero tensor. For details, see [14, 15].

In discretized space, a perfectly matched interface does not exist [6], and hence, reflections do occur in computations. Fortunately, they can be made to be negligible. In our implementation, this is done by choosing the depth of the PMLs carefully.

We also need to model the PMLs found in regions where $x \in X_- \cup X_+$. Using arguments equivalent to those needed to derive (24)-(26), the following system of equations is obtained:

$$j\omega \left(1 + \frac{\hat{\sigma}(x)}{j\omega\epsilon_0}\right) \left(1 + \frac{\hat{\sigma}(z)}{j\omega\epsilon_0}\right) \widehat{D}_y = c_0 \left(\frac{\partial \widehat{H}_z}{\partial x} - \frac{\partial \widehat{H}_x}{\partial z}\right) \tag{28}$$

$$j\omega \left(1 + \frac{\hat{\sigma}(x)}{j\omega\epsilon_0}\right)^{-1} \left(1 + \frac{\hat{\sigma}(z)}{j\omega\epsilon_0}\right) \widehat{H}_x = c_0 \frac{\partial \widehat{E}_y}{\partial z} \tag{29}$$

$$j\omega \left(1 + \frac{\hat{\sigma}(x)}{j\omega\epsilon_0}\right) \left(1 + \frac{\hat{\sigma}(z)}{j\omega\epsilon_0}\right)^{-1} \widehat{H}_z = -c_0 \frac{\partial \widehat{E}_y}{\partial x}. \tag{30}$$

This system together with (12) gives us the full system of equations that we wish to solve. (For a thorough overview of PML technology for electromagnetics problems see [16] and [8].)

We now must write (28)-(30) in the time domain.

2.3. The Full Formulation in the Time Domain

It is a straightforward exercise to show that (28)-(30) can be expressed in the time domain as

$$\partial_t D_y^* + \frac{\hat{\sigma}(x)}{\epsilon_0} D_y^* = c_0 \left(\partial_x H_z - \partial_z H_x\right), \tag{31}$$

$$\partial_t D_y + \frac{\hat{\sigma}(z)}{\epsilon_0} D_y = \partial_t D_y^*, \tag{32}$$

$$\partial_t H_x^* + \frac{\hat{\sigma}(z)}{\epsilon_0} H_x^* = c_0 \partial_z E_y, \tag{33}$$

$$\partial_t H_x = \partial_t H_x^* + \frac{\hat{\sigma}(x)}{\epsilon_0} H_x^*, \tag{34}$$

$$\partial_t H_z^* + \frac{\hat{\sigma}(x)}{\epsilon_0} H_z^* = -c_0 \partial_x E_y, \tag{35}$$

$$\partial_t H_z = \partial_t H_z^* + \frac{\hat{\sigma}(z)}{\epsilon_0} H_z^*, \tag{36}$$

where D_y^*, H_x^*, and H_z^* are defined by equations (32), (34), and (36) respectively. Note that outside of the PML regions, $\hat{\sigma}(x) = \hat{\sigma}(z) = 0$ and (31)-(36) reduces to (19)-(21). (Recall that we have suppressed the "~" notation.)

The y component of the time domain expression of (12) is given by

$$D_y(t) = \epsilon_\infty E_y(t) + P_y(t) + C_y(t), \tag{37}$$

where the polarization P_y satisfies

$$\tau \dot{P}_y + P_y = (\epsilon_s - \epsilon_\infty) E_y, \tag{38}$$

and the conductivity term C_y satisfies

$$\dot{C}_y = (\sigma/\epsilon_0) E_y \tag{39}$$

inside of the dielectric. Outside of the dielectric, $C_y = P_y = 0$. Recall that we rescaled E_y via equation (10) and then suppressed the corresponding notation change. This accounts for the change in form from (8) to (38). Also, note that D_y in (37) is *not* the y component of the displacement vector $\mathbf{D}$ of equation (7). It is, rather, the y component of $\widehat{\widetilde{\mathbf{D}}}$ in (12) written in the time domain. The "~" was suppressed at the beginning of Section 2.2 in order to simplify notation.

Equations (31)-(38) are the full set of equations that we solve numerically. We use finite differences. In particular, we use the FDTD algorithm of [15, 16], which uses forward differences to approximate the time and spacial derivatives. The Yee staggered grid is used [15, 16], which ensures second order accuracy within the free space region. Unfortunately, this accuracy is not maintained within or at the boundary of the dielectric medium. The FDTD algorithm is fully explicit, and does not require the inversion of large, sparse matrices. Our implementation of the FDTD algorithm for this problem, including the implementation of the PMLs, follows details found in [15] closely, with the exception that the corner regions in which there is overlap of PML layers are dealt with as in [7]. Gauss' divergence laws (1) and (2) are automatically enforced by the FDTD algorithm (see [16]). Well-posedness for this system is an important question, which we hope to address in a future paper.

In all simulations, the spatial and time steps are related via

$$\Delta t = \frac{\Delta x}{2c_0}, \tag{40}$$

where $c_0(= 3 \times 10^8)$ is the (approximate) speed of light. This satisfies the CFL condition, and hence, guarantees the stability of the FDTD algorithm in free space [16]. Fortunately, stability is retained when a Debye medium is introduced [12].

An unfortunate side effect of numerical solutions to Maxwell's equations is that dispersion is introduced by the numerical scheme. This is the case for both finite difference and finite element (FEM) methods (see [8, 16]). Techniques can be used to minimize dispersion but such techniques are very difficult to implement when dispersive materials are present [12]. For the Debye medium in particular, it is noted in [12] that when the FDTD algorithm is used spurious dispersion may occur unless $\Delta t \sim \mathcal{O}(10^{-3}\tau)$, which can be a highly restrictive condition.

A different, but standard, approach used in solving Maxwell's equations in the time domain is to first construct the wave equation for the electric field **E**, and then solve it numerically. In [2] this is done for the forward problem addressed in this paper using FEM. The main drawback of the FEM approach is that at each time step a large sparse linear system must be solved. This is not the case for the FDTD algorithm. Consequently, for the same grid spacing, FDTD is much faster. In addition, the implementation of the FDTD algorithm is much simpler for this problem than it is for FEM. This is particularly true for the implementation of the PMLs. In the numerical comparisons that we have done between these two approaches, the FDTD algorithm was at least ten times faster when the same computational grid was used. On the other hand, we have not carried out a convergence comparison between FDTD and FEM for the forward problem. It is likely that the forward solutions obtained using FEM will converge on a coarser grid than does FDTD, since FDTD loses its second order accuracy within the dielectric and at the dielectric/free space interface. If this is the case, savings in computational time may be lost. Hence, it is not clear at this point which approach (FEM or FDTD) is better for use in solving the inverse problem. These questions are currently being addressed in further efforts.

2.4. The Forward Problem for a Debye Medium

In this section we present details and results from an attempt at solving (31)-(38) using the FDTD approach discussed in the previous section. We begin with the geometry of the computational region. The necessary measurements are given by $X_0 = [0, 0.1]$, $Z_v = [0, 0.2]$, $Z_d = [0.2, 0.4]$ (these measurements are in meters). In order to achieve convergence of the finite difference approximations we take the number of nodes along the z-axis to be $N = 640$. The spatial step size in both the x and z directions is given by

$$\Delta x = 0.4/N. \tag{41}$$

Then, from (40) we obtain $\Delta t \approx 1.0417 \times 10^{-12}$ seconds.

The parameters that give the position of the source term J_s defined in (9) are $\bar{x} = 0.025$ and $\bar{z} = 0.1$. For our simulations, the frequency of our sine wave is 3 GHz, and hence $\hat{\omega} = 2\pi * 3 \times 10^9$. The duration t_f of the sine pulse is 0.667 nanoseconds (ns), which is two complete periods.

The depth of the PML layers is 0.05 meters. This measurement balances computational efficiency with the minimization of numerical noise caused by the reflections from the PML interfaces. The spatial discretization within the PML is the same as that within the computational region.

In order to attempt the parameter identification problem, we must collect data within the computational domain at a pre-specified sampling rate. We collect data at the center of the antenna at each time step. Then, our data consists of the set $\{E_y(i\Delta t, 0, \bar{z}, \mathbf{q})\}_{i=1}^{N_t}$, where E_y is the solution of Maxwell's equations given by the FDTD algorithm, $\mathbf{q} = (\sigma, \tau, \epsilon_s, \epsilon_\infty)$ is the set of parameters that characterize the Debye medium,
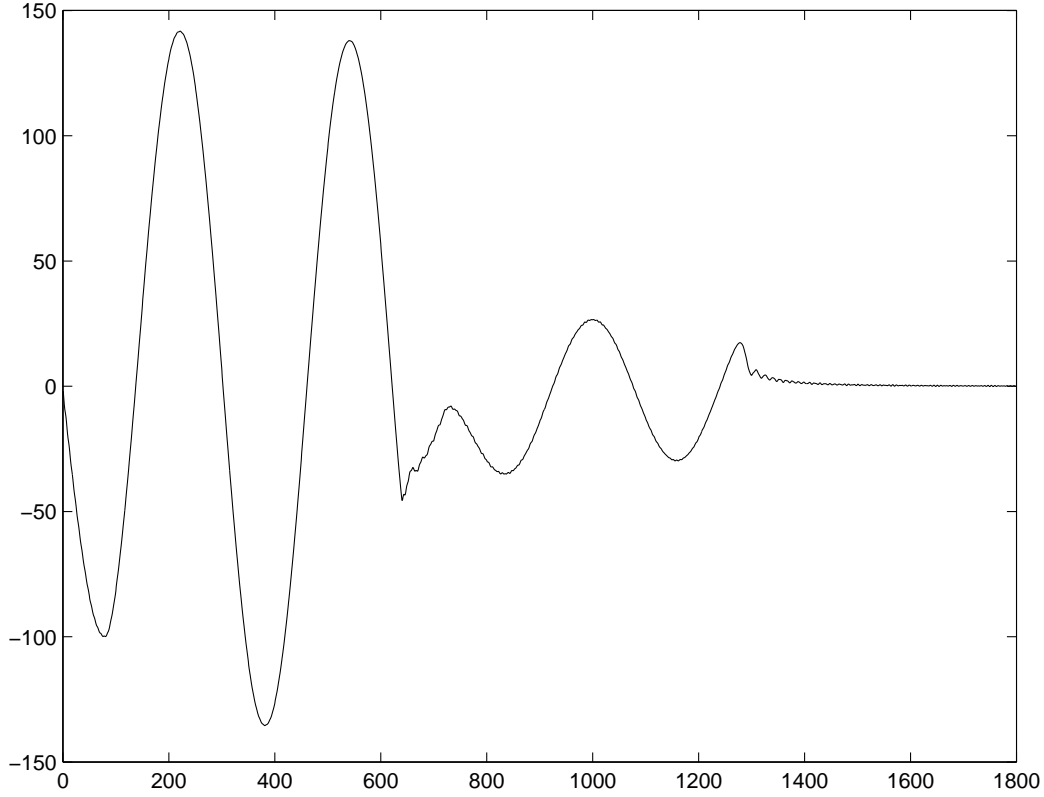
Figure 3. Data for the Debye model for water. The vertical axis gives the magnitude of the electric field. The horizontal axis gives the time in the units Δt seconds.

and N_t is the total number of time steps. Notice that we have now made explicit the dependence of solutions of (31)-(38) on the parameter set $\mathbf{q}$. We will continue to use this convention in the sequel. For the discretization mentioned above, we take $N_t = 1800$, in which case the data includes both the outgoing energy and the energy that is reflected from the free-space/dielectric interface.

In [5] it is noted that the polarization behavior of water is reasonably modelled by (37), (39), and (38) with parameter values:

$$
\begin{aligned}
\sigma &= 1 \times 10^{-5} \quad \text{mhos/meter,} \\
\tau &= 8.1 \times 10^{-12} \quad \text{seconds,} \\
\epsilon_s &= 80.1 \quad \text{relative static permittivity,} \\
\epsilon_\infty &= 5.5 \quad \text{relative high frequency permittivity.}
\end{aligned}
\tag{42}
$$

With this information, numerical approximations to (31)-(38) can be obtained. The resulting data set $\{E_y(i\Delta t, 0, \bar{z}, \mathbf{q})\}_{i=1}^{N_t}$ is plotted in Figure 3. In Figure 4, we see the formation of the Brillouin precursors [1, 2, 3] within the dielectric. This suggests that the FDTD solutions are accurately capturing the dispersivity of the Debye medium.
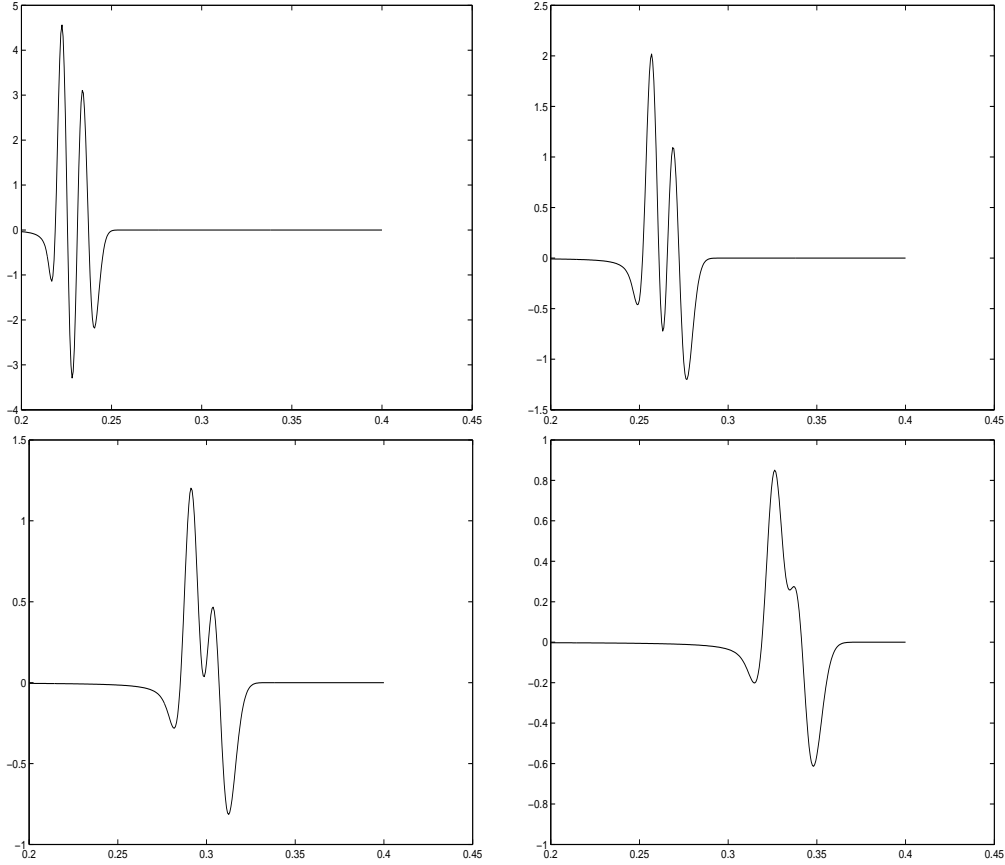
Figure 4. Precursor Formation Within the Debye Medium. We plot $\{E_y(\bar{t}, 0, z)\}$, for $z \in [0.2, 0.4]$ at four different times. In the upper left, $\bar{t} = 1500\Delta t$. In the upper right, $\bar{t} = 2500\Delta t$. In the lower left, $\bar{t} = 3500$. In the lower right, $\bar{t} = 4500$. These times correspond to 1.56, 2.60, 3.65, and 4.69 ns respectively.

3. The Inverse Problem

We now address the problem of parameter identification for a Debye medium. In the previous section, the problem of solving Maxwell's equations with a Debye polarization model was addressed. Implicit in this problem is the knowledge of the "true" parameter vector $\mathbf{q}^*$, given by (42). In the inverse problem, on the other hand, we have measurements of the electric field at the point $(0, \bar{z})$ in the computational domain at uniform intervals of time t_i. Given this data, our goal is to then identify the "true" parameter vector $\mathbf{q}^*$.

3.1. The Noise Model for Data Generation

In the laboratory setting, noise enters into the problem both through the resistor, which generates the electric pulse, and through the antenna/receiver, which records the electric field intensities at the point $(0, \bar{z})$ at discrete time steps. The noise model presented in this section follows [13] closely.

First, noise enters the signal via the random motion of electrons in the resistor.

These random motions produce small, random voltage fluctuations at the resistor terminals. Over time these voltage fluctuations have a zero average value, but non zero rms (or standard deviation) given, at time t_i, by Planck's black body radiation law,

$$v_i = \sqrt{\frac{4hfBR}{e^{hf/KT} - 1}}, \tag{43}$$

where

$$h = 6.546 \times 10^{-34} \text{ J-sec is Planck's constant,}$$
$$k = 1.380 \times 10^{-23} \text{ J/}^o\text{K is Boltzmann's constant,}$$
$$T \text{ is the temperature is degrees kelvin (K),}$$
$$B \text{ is the bandwidth of the system in Hz,}$$
$$f \text{ is the center frequency of the bandwidth in Hz,}$$
$$R \text{ is the resistance.}$$

For microwave frequencies we have $hf << kT$, and hence,

$$e^{hf/KT} - 1 \approx \frac{hf}{KT}.$$

Thus equation (43) is well approximated by

$$v_i = \sqrt{4kTBR}. \tag{44}$$

This is the approximation commonly used in microwave work [13]. The noise is modelled as independent and identically distributed (i.i.d.) Gaussian with constant variance v_i^2.

If we replace the noisy resistor with a noiseless resistor together with a voltage generator [13], then we have

$$v_i = \sqrt{kTB}. \tag{45}$$

Noise also enters the data through the antenna and receiver. The noise comes from two sources: the noise due to the antenna, and the electronic noise in the receiver. Both noises are assumed to be white Gaussian with standard deviation $v_i = \sqrt{kTB}$ [10, 13].

With this model, the noise in the data is zero mean, additive Gaussian, with constant variance across time. That is, our data vector

$$\mathbf{E}_y^{data} = (E_{y,1}^{data}, \ldots, E_{y,N_t}^{data}) \tag{46}$$

is a random vector with normally distributed components

$$E_{y,i}^{data} \sim \mathcal{N}(E_y(t_i, 0, \bar{z}, \mathbf{q}^*), \Sigma^2) \quad \text{for} \quad i = 1, \ldots, N_t, \tag{47}$$

where Σ^2 is calculated using the above approximations for the variances of the noise that comes from the resistor, the antenna, and the receiver.

To generate synthetic, noisy data, we use the FDTD method, with the specifics of Section 2.4, to obtain the set $\{E_y(t_i, 0, \bar{z}, \mathbf{q}^*)\}_{i=1}^{N_t}$, together with noise model (47). We then seek the maximum likelihood estimator (MLE) of the true parameter vector $\mathbf{q}^*$

given by (42). Given the noise model (47), with data equal to n realizations $\{\mathbf{d}_j\}_{j=1}^n$ of (46),(47), the MLE $\hat{\mathbf{q}}_n$ for $\mathbf{q}^*$ is given by

$$\hat{\mathbf{q}}_n = \arg\min_{\mathbf{q} \in Q} J(\mathbf{q}), \qquad \text{where} \qquad J(\mathbf{q}) = \frac{1}{2} \sum_{j=1}^n ||\mathbf{E}_y(0, \bar{z}, \mathbf{q}) - \mathbf{d}_j||_2^2, \quad (48)$$

where $\mathbf{E}_y(0, \bar{z}, \mathbf{q}) = (E_y(t_1, 0, \bar{z}, \mathbf{q}), \ldots, E_y(t_{N_t}, 0, \bar{z}, \mathbf{q}))$, and Q is the constraint set for the parameter values.

3.2. The Optimization Problem

The minimization problem (48) is solved using a modified Levenberg-Marquardt method [11]. In the particular implementation that we employ, an ellipsoidal trust region approach is taken, which alleviates difficulties caused by large differences in the magnitudes of the parameters in the vector $\mathbf{q}$. In particular, each variable is scaled to be of roughly the same magnitude, namely each scaled variable is approximately $\mathcal{O}(1)$ near $\mathbf{q}^*$. Prior knowledge of the magnitudes of the parameters at the solution can be used to determine the scaling factors. The trust-region subproblem is solved at each outer iteration using the dogleg method (see [11]). In addition, we enforce the physical constraints $\sigma \geq 0$, $\tau \geq 0$, $\epsilon_s \geq 1$, and $\epsilon_\infty \geq 1$, which determine Q. This is done via the projection of both the trust region and of the search directions onto the feasible set at each iteration. The convergence properties of the unconstrained version of the algorithm are then retained. The details of this algorithm will be presented in a future paper.

The Jacobian matrix, which is used both for the computation of the gradient and for the construction of the Gauss-Newton approximation to the Hessian, is approximated using forward differences. Each iteration then involves at most five function evaluations: one at the current iterate, and one for each of the four parameters. This is not overly restrictive for our problem, but for problems in which the permittivity and conductivity are allowed to be spatially varying, this is no longer a viable approach. Analytic computations of the gradient are possible, would alleviate this difficulty, and require no finite difference approximations. We hope to explore the use of analytic gradients for this problem in future work.

It is important to note that a solution to (48) depends upon the spatial discretization Δx given by (41) (note that Δt is then obtained using (40)), which depends only on the number of grid points N in $Z_v \cup Z_d$ of Figure 2. (We assume that the length of 0.4 for $Z_v \cup Z_d$ is fixed.) Thus in (48), J and any approximate solution q of (48) should carry the index N. Nonetheless, in the sequel we will suppress this N dependence.

3.3. The Inverse Problem for a Debye Medium

We now present results for the inverse problem of identifying the dielectric parameters in the Debye model for water given noisy data. Using the parameters given in (42) and the discretization details given in Section 2.4, we generate data with varying levels

Table 1. Estimated parameters in Debye inverse problem. Test 1

Test	% Noise	σ	τ	ϵ_s	ϵ_∞	Residual
True values:		1.0×10^{-5}	8.1×10^{-12}	80.1	5.5	
1	0.0	0.0	9.965×10^{-12}	80.75	14.77	6.7051×10^{-8}
2	0.5	0.0	9.974×10^{-12}	80.20	13.30	1.2279×10^{-5}
3	1.0	0.0	9.980×10^{-12}	80.18	12.23	4.8912×10^{-5}
4	1.5	1.952×10^{-1}	7.437×10^{-12}	79.07	10.00	1.1483×10^{-4}
5	2.0	0.0	9.989×10^{-12}	80.01	10.66	1.9542×10^{-4}
6	2.5	3.671×10^{-2}	9.997×10^{-12}	79.57	9.717	3.1861×10^{-4}
7	3.0	8.715×10^{-2}	9.998×10^{-12}	79.82	9.189	4.5844×10^{-4}
8	5.0	0.0	9.903×10^{-12}	85.42	16.22	1.2000×10^{-3}

of noise using the error model (47) with $t_i = i\Delta t$. Hence, our sampling rate for the antenna/receiver is assumed to be $\Delta t \approx 0.001$ ns. Admittedly, such a sampling rate may not be possible in many laboratory settings. Several numerical experiments were performed in which data was generated using error model (47) with $t_i = i \cdot \ell \cdot \Delta t$, where ℓ is an integer larger than 1. As one would expect, our ability to identify the true parameter vector $\mathbf{q}^*$ in the inverse problem decreased as ℓ increased. Once real data is obtained, with a particular sampling rate, it will likely be necessary to decrease the sampling rate when solving the inverse problem.

We make eight different attempts at the inverse problem with a different noise level for each attempt. We solve (48) using the approach discussed in the previous section with initial guess $\mathbf{q}_0 = (\sigma_0, \tau_0, \epsilon_{s,0}, \epsilon_{\infty,0})$ given by $\sigma_0 = 1.5 \times 10^{-5}$, $\tau_0 = 10.0 \times 10^{-12}$, $\epsilon_{s,0} = 73.1$, and $\epsilon_{\infty,0} = 6.0$. (These range between 50% relative to 10% relative error from the true values and are typical of values used in testing algorithms [4].). We use only one data vector $\mathbf{d}_1$. Then $n = 1$ in (48). The results are given in Table 1. The residual, which is given in the far right column, is $J(\hat{\mathbf{q}})/\|\mathbf{d}_1\|^2$, where $\hat{\mathbf{q}}$ is the approximate solution to (48) given by the optimization algorithm.

A partial explanation for results in Table 1 can be obtained from an analysis of equation (14), which we rewrite as

$$\epsilon_r^*(\omega) = \frac{\epsilon_s}{1 + j\omega\tau} + \frac{j\omega\tau}{1 + j\omega\tau}\epsilon_\infty + \frac{\sigma}{j\omega\epsilon_0}. \tag{49}$$

Since, in our simulations, the outgoing radiation is given by two cycles of a sine wave, in the continuum space, the resulting pulse will have infinite frequency content. Nonetheless, for the purpose of the inverse problem, the outgoing and reflected radiation will be dominated by frequencies near the carrier frequency, 3 GHz. This is seen in Figure 5. Thus ϵ_r^* will be dominated by frequencies near 3 GHz, and so we restrict our analysis of (49) to the case where $\omega \approx 3 \times 10^9$. Then, since $\omega\tau^* \sim \mathcal{O}(10^{-2})$,

$$\frac{\epsilon_s^*}{1 + j\omega\tau^*} \sim \mathcal{O}(\epsilon_s^*), \tag{50}$$
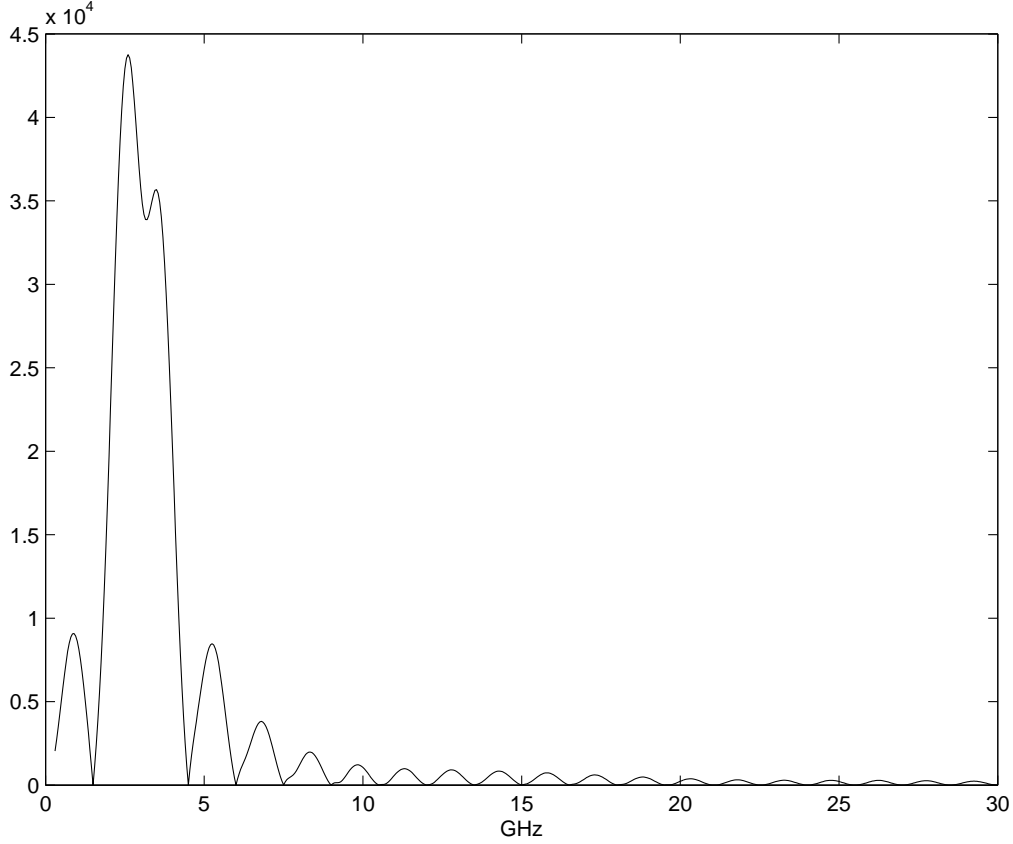
Figure 5. Fourier Transform of the Data in Figure 3. The vertical axis gives amplitude. The horizontal axis gives the frequency from 0.3 GHz to 30 GHz.

$$\frac{j\omega\tau^*}{1 + j\omega\tau^*}\epsilon_\infty^* \sim \mathcal{O}(10^{-2}\epsilon_\infty^*). \tag{51}$$

Furthermore, given that $\epsilon_0 \approx 8 \times 10^{-12}$, we have

$$\frac{\sigma^*}{j\omega\epsilon_0} \sim \mathcal{O}(10^2\sigma^*). \tag{52}$$

Equations (50), (51), and (52) suggest that ϵ_r^* will be most sensitive to the static permittivity ϵ_s. More specifically, near $\mathbf{q}^*$, ϵ_r^* will is more sensitive to changes in ϵ_s than to changes in τ, ϵ_∞, and σ of the same relative magnitude. Further support for this claim can be found by an appeal to the reflection coefficient Γ, which is defined by

$$\Gamma = \frac{1 - \sqrt{\epsilon_r^*}}{1 + \sqrt{\epsilon_r^*}}.$$

Γ will also be dominated by ϵ_s. We therefore expect that the cost function J will be most sensitive to ϵ_s, and hence, that ϵ_s will be the easiest parameter to reconstruct. This is supported by the results in Table 1.

Furthermore, equations (50) and (51) suggest that near τ^*, ϵ_r^* is relatively insensitive to "small" changes in τ, since if $\tau \approx \tau^*$, the approximation $1 + j\omega\tau \approx 1$ remains valid. On the other hand, if τ becomes too large, it's effects will be manifested in both the first

and second terms of (49). In fact, we found that if τ is fixed and is large in comparison to τ^*, the reconstructed values of ϵ_s are very poor. These observations explain why in Table 1 the reconstructed values of τ are near in magnitude to τ^* at all noise levels.

Since ϵ_r^* is dominated by ϵ_s, we would not expect the accurate reconstruction of ϵ_∞^*. Furthermore, assuming

$$\mathcal{O}(\tau) = \mathcal{O}(\tau^*), \tag{53}$$

$$\tau\epsilon_\infty \approx \tau^*\epsilon_\infty^*, \tag{54}$$

from equation (51), we have

$$\frac{j\omega\tau}{1+j\omega\tau}\epsilon_\infty \approx \frac{j\omega\tau^*}{1+j\omega\tau^*}\epsilon_\infty^*. \tag{55}$$

Hence, independent identification of τ^* and ϵ_∞^* will be very difficult, particularly in the presence of noise. The identification of their product, on the other hand, will be less difficult, though we would not expect to reconstruct $\tau^*\epsilon_\infty^*$ as accurately as we do ϵ_s^*. This is also supported by the results of Table 1.

The observations of the previous two paragraphs suggest a strategy for the inverse problem. Namely, fix the value of ϵ_∞ to an approximate of ϵ_∞^* of the same order and optimize over the other parameters. The optimized value of τ will then likely satisfy (53), (54), and (55), in which case the dynamics of the problem will be adequately captured. Before testing this approach, we discuss the parameter σ.

Equation (52) suggests that J is least sensitive to the parameter σ. In addition, since the change in σ from its value in free space to its value in the dielectric is very small, we would expect very little of the reflected energy to be due to the change in σ. The results of Table 1 are therefore no surprise. Also not surprising is the fact that the magnitude of σ remains small in all reconstructions. This suggests the strategy of fixing the value of σ to a value reasonably close to σ^* prior to optimization. The dynamics of the problem should then be adequately captured.

We now test the strategies suggested in the previous two paragraphs. Namely, we perform the inverse problem with the values of σ and ϵ_∞ fixed at $\sigma = 0.0$ and $\epsilon_\infty = 1.0$. The results of this approach are presented in Table 2. The values of $\epsilon_{s,0}$ and τ_0 are the same as in the previous attempt. Comparing these results with those of Table 1, we see that this strategy yields very little change in the values of the residuals, while the reconstruction of ϵ_s remains accurate. In addition, the solution to the inverse problem can be had far more quickly when σ and ϵ_∞ are fixed at the outset. Though more iterations are needed in this case, gradient computations are obtained with 3/5ths the cost. Also, once σ and ϵ_∞ are fixed, constraints are no longer necessary.

The relative stability of the τ reconstruction in Table 2 suggests yet another strategy. Namely, fix the value of τ at the outset. Results for this approach are given in Table 3 where τ was fixed to be the initial guess τ_0 from the previous attempts. Surprisingly, we find the this strategy actually yields slightly smaller values for the residuals than were obtained in Table 2. In addition, the optimization problems converged far more quickly. In fact, at all noise levels, within three iterations our

Table 2. Estimated parameters in Debye inverse problem. Test 2.

Test	% Noise	σ	τ	ϵ_s	ϵ_∞	Residual
True values:		1.0×10^{-5}	8.1×10^{-12}	80.1	5.5	
1	0.0	0.0	9.818×10^{-12}	80.82	1.0	5.8619×10^{-7}
2	0.5	0.0	9.767×10^{-12}	80.42	1.0	1.3168×10^{-5}
3	1.0	0.0	9.781×10^{-12}	80.94	1.0	5.2418×10^{-5}
4	1.5	0.0	9.751×10^{-12}	82.09	1.0	1.1127×10^{-4}
5	2.0	0.0	9.922×10^{-12}	82.69	1.0	1.9741×10^{-4}
6	2.5	0.0	9.123×10^{-12}	79.82	1.0	3.1841×10^{-4}
7	3.0	0.0	9.598×10^{-12}	79.44	1.0	4.5081×10^{-4}
8	5.0	0.0	9.703×10^{-12}	79.51	1.0	1.2000×10^{-3}

Table 3. Estimated parameters in Debye inverse problem. Test 3.

Test	% Noise	σ	τ	ϵ_s	ϵ_∞	Residual
True values:		1.0×10^{-5}	8.1×10^{-12}	80.1	5.5	
1	0.0	0.0	1.000×10^{-11}	80.75	1.0	6.8509×10^{-7}
2	0.5	0.0	1.000×10^{-11}	81.24	1.0	1.2546×10^{-5}
3	1.0	0.0	1.000×10^{-11}	80.03	1.0	5.0736×10^{-5}
4	1.5	0.0	1.000×10^{-11}	80.26	1.0	1.1028×10^{-4}
5	2.0	0.0	1.000×10^{-11}	82.46	1.0	1.9746×10^{-4}
6	2.5	0.0	1.000×10^{-11}	80.11	1.0	3.1851×10^{-4}
7	3.0	0.0	1.000×10^{-11}	81.74	1.0	4.4774×10^{-4}
8	5.0	0.0	1.000×10^{-11}	78.72	1.0	1.2000×10^{-3}

optimization algorithm had converged. Using the previous strategies, many more iterations were required.

The previous approach rested heavily on the fact that our choice of τ_0 was near in magnitude to τ^*. An obvious question, then, is what can be done if one cannot obtain a good initial guess for τ^*? We take the approach of fixing the parameters $\sigma = 0$, $\epsilon_s = 73.1$, and $\epsilon_\infty = 1$ and optimizing, with no noise, with respect to the parameter τ. With an initial guess of $\tau_0 = 10^{-4}$, we obtain the approximation $\tau = 6.9987 \times 10^{-12}$. We observe how close this is to τ^*. We then fix this value of τ and optimize over ϵ_s. The results of this approach are given in Table 4 and are similar to those found in Table 3.

3.4. Reconstructions with Inferior Accuracy

In each of the above experiments, the data was generated on the same grid that was used in the inverse problem, and i.i.d. Gaussian noise was added. In this section, we attempt the inverse problem with data generated instead on a finer grid. No additional

Table 4. Estimated parameters in Debye inverse problem. Test 4.

Test	% Noise	σ	τ	ϵ_s	ϵ_∞	Residual
True values:		1.0×10^{-5}	8.1×10^{-12}	80.1	5.5	
1	0.0	0.0	6.9987×10^{-12}	79.82	1.0	5.7768×10^{-8}
2	0.5	0.0	6.9987×10^{-12}	79.67	1.0	1.2299×10^{-5}
3	1.0	0.0	6.9987×10^{-12}	80.18	1.0	5.4274×10^{-5}
4	1.5	0.0	6.9987×10^{-12}	80.02	1.0	1.1609×10^{-4}
5	2.0	0.0	6.9987×10^{-12}	78.54	1.0	2.0476×10^{-4}
6	2.5	0.0	6.9987×10^{-12}	78.64	1.0	3.1835×10^{-4}
7	3.0	0.0	6.9987×10^{-12}	82.33	1.0	4.4185×10^{-4}
8	5.0	0.0	6.9987×10^{-12}	78.67	1.0	1.3000×10^{-3}

Table 5. Estimated parameters in Debye inverse problem. Test 5.

Test	σ	τ	ϵ_s	ϵ_∞	Residual
True values:	1.0×10^{-5}	8.1×10^{-12}	80.1	5.5	
1	0.0	6.9987×10^{-12}	72.22	1.0	3.8477×10^{-5}
2	1.0×10^{-5}	8.1000×10^{-12}	72.55	5.5	3.6494×10^{-5}

noise is added. More specifically, we generate our data on the grid with $N = 1280$ in (41) and attempt the inverse problem on the grid with $N = 640$. In our first experiment, we attempt the inverse problem with σ, τ, and ϵ_∞ fixed to the values found in Table 4. These results are given by Test 1 in Table 5. The reconstructed value of ϵ_s is disappointingly far from ϵ_s^*. In order to see if the choice of fixed parameters is the culprit, we solve the inverse problem with the σ, τ, and ϵ_∞ fixed at the true values σ^*, τ^*, and ϵ_∞^* respectively. There is very little difference in the result, which is given by Test 2 in Table 5.

One possible explanation for these results is that the FDTD algorithm has not quite converged for $N = 640$, even though plots of $\{E_y(t_i, 0, \bar{z})\}$, such as is found for $N = 640$ in Figure 3, seem to suggest that convergence has occurred. Unfortunately, solving the inverse problem on the grid defined by $N = 1280$ in (41) is computationally challenging.

If the FDTD algorithm has not converged for $N = 640$, then one is led to question the results and the conclusions of Section 3.3. Fortunately, numerical experiments suggest that they will remain unchanged for $N = 1280$. Also, for $N = 320$ and $N = 160$, the same experiments were performed with very similar results and conclusions. Moreover, the results and conclusions of Section 3.3 are very similar to those found in [3] for the analogous one dimensional problem using FEM. Hence, we conclude that they are valid. However, the above comments do provide a caution that the values of parameter estimates may depend significantly on resolution of the underlying dynamical

systems.

One should also question the strategy of using FDTD when solving the inverse problem. The FDTD algorithm, though second order accurate in free space, looses that accuracy within the PMLs and, more importantly, within the dielectric [12]. Furthermore, FDTD is highly vulnerable to error at the free space/dielectric interface. Thus, though FDTD is unmatched in simplicity of implementation for this problem, it may not be the proper forward solver to use if the goal is the solution of the inverse problem. Clearly, further investigation is necessary before this question and similar ones are resolved.

4. Estimating Variability in Reconstructions

In the previous section, we obtain reconstructions of the true parameter vector $\mathbf{q}^*$ by approximately solving (48) when $n = 1$. In each of the above approaches, the cost function J in depends upon a data vector $\mathbf{d}_1$, which is a realization of the random vector (46),(47). If the experiment that was performed to obtain $\mathbf{d}_1$ is performed a second time, a different data vector $\mathbf{d}_2$ is obtained, which is also a realization of (46),(47). If the computations in the previous section are repeated using $\mathbf{d}_2$ in place of $\mathbf{d}_1$, the cost function J will be different, and hence, different values for the reconstructed parameter vector $\hat{\mathbf{q}}$ will be obtained. This illustrates the fact that the reconstruction $\hat{\mathbf{q}}(\mathbf{d})$ is in fact random as well. A natural question therefore arises. Namely, what sort of variability is in our reconstruction? Or, put in another way, what kind of confidence can we have that our reconstruction $\hat{\mathbf{q}}$ is close to the true parameter vector $\mathbf{q}^*$?

To address this question, we will make use of large-sample theory from statistics (see Chapter 7 in [9]). In addition, we will restrict our attention to the case of Table 4, where the parameters σ, τ, and ϵ_∞ are fixed. Then the cost function J in (48) depends only on ϵ_s. We now state the theorem from [9], p. 469, that we will use.

Theorem 3.1: *Let $\mathbf{D}_1, \ldots, \mathbf{D}_n$ be independent and identically distributed random vectors that depend on a parameter ϵ_s and have probability density function f_{ϵ_s}. If f_{ϵ_s} satisfies certain regularity conditions, then any consistent sequence $\hat{\epsilon}_{sn} = \hat{\epsilon}_{sn}(\mathbf{D}_1, \ldots, \mathbf{D}_n)$ of roots of the likelihood equation satisfies*

$$\sqrt{n}(\hat{\epsilon}_{sn} - \epsilon_s^*) \to \mathcal{N}(0, 1/I(\epsilon_s^*)), \tag{56}$$

where ϵ_s^* is the "true" parameter value, and it is assumed that

$$I(\epsilon_s^*) = \mathcal{E}_{\epsilon_s}\left[-\frac{\partial^2}{\partial \epsilon_s^2} \log f_{\epsilon_s^*} \right] \tag{57}$$

satisfies $0 < I(\epsilon_s^*) < \infty$.

The random vectors $\mathbf{D}_1, \ldots, \mathbf{D}_n$ in Theorem 3.1 are given by

$$\mathbf{D}_i = \mathbf{E}_y^{data} \qquad i = 1, \ldots, n,$$

where $\mathbf{E}_{data}$ is defined by (46), (47), and are therefore independent and identically distributed.

The density function f_{ϵ_s} is related to the cost function J in (48) via

$$J(\epsilon_s) = -\Sigma^2 \cdot \log f_{\epsilon_s} + c, \tag{58}$$

where c is a constant. Hence,

$$I(\epsilon_s^*) = \mathcal{E}_{\epsilon_s} \left[\frac{1}{\Sigma^2} \cdot \frac{\partial^2}{\partial \epsilon_s^2} J(\epsilon_s^*) \right]. \tag{59}$$

In order to use Theorem 3.1, we must assume that the cost function J has a bounded third derivative. This amounts to the assumption that the mapping from ϵ_s to $E_y(t, 0, \bar{z}, \epsilon_s)$ has a bounded third derivative, which will be true almost everywhere. The remaining regularity conditions required for f_{ϵ_s} by Theorem 3.1 are easily satisfied.

Numerical experiments suggest that $J(\epsilon_s)$ is strictly convex, in which case the only local minimum of J is the global minimum. Then, if our algorithm creates a sequence for which the derivative of J converges to zero, this sequence must converge to the maximum likelihood estimator. Given that this occurs for each n, the sequence $\{\hat{\epsilon}_{s,n}\}_{n=1}^{\infty}$ will be consistent.

From (56) we know that

$$\hat{\epsilon}_{s,n} \rightarrow \epsilon_s^*.$$

We define

$$\alpha_n = \frac{1}{\sigma^2 n} J''(\hat{\epsilon}_{s,n}). \tag{60}$$

It is shown in [9] that

$$\alpha_n - I(\hat{\epsilon}_{s,n}) \rightarrow 0,$$

and hence

$$\alpha_n \rightarrow I(\epsilon_s^*). \tag{61}$$

We will therefore use α_n to approximate $I(\epsilon_s^*)$ in (57).

In Table 6 we solve problem (48) for various values of n at the 2% noise level. We note that as n increases, the values for $\hat{\epsilon}_{s,n}$ and α_n converge, then appear to begin to diverge at $n = 64$. The divergence is likely due to the fact that for large values of n the finite difference approximations to the gradient and Hessian become inaccurate for the function J in (48).

The rapid convergence of $\{\alpha_n\}$ suggests that the approximations of $I(\epsilon_s^*)$ are reasonably accurate for small values of n, and hence, that the values obtained for the approximate standard deviation, which are found in the last column of Table 6, should be reasonably accurate as well. This is supported by the results found in Table 6, where we see that for each n, with the exception of $n = 1$ and $n = 64$, $\hat{\epsilon}_{s,n}$ is within one standard deviation of the true value ϵ_s^*. For $n = 1$ and $n = 64$, the reconstruction is within two standard deviations of ϵ_s^*, which is also reasonable. We note also that, as predicted by Theorem 3.1, $\hat{\epsilon}_{s,n}$ converges, more or less, to the true value $\epsilon_s^* = 80.1$.

Table 6. Estimates for the Variability in $\hat{\epsilon}_s$ for the Example with 2% Noise. The values of σ, τ, and ϵ_∞ are fixed at the values found in Figure 4. We use α_n to estimate $I(\epsilon_s^*)$, $1/n\alpha_n$ to estimate the variance, and $1/\sqrt{n\alpha_n}$ to estimate the standard deviation.

n	$\hat{\epsilon}_{s,n}$	α_n	$1/n\alpha_n$	$1/\sqrt{n\alpha_n}$
1	82.07	0.3906	2.5604	1.6001
2	80.36	0.4146	1.2060	1.0982
4	79.84	0.4211	0.5937	0.7705
8	80.38	0.4138	0.3021	0.5496
16	80.11	0.4177	0.1496	0.3868
32	80.01	0.4194	0.0745	0.2729
64	79.73	0.4233	0.0369	0.1921

Finally, we note that for $n = 1$, the approximation to $I(\epsilon_s^*)$ is near to the converged value. Also, the value of $\hat{\epsilon}_{s,1}$ is reasonably close to the true value ϵ_s^*. Hence, the corresponding approximate to the standard deviation may give a reasonable approximation of the true variability in the reconstruction $\hat{\epsilon}_{s,1}$. In the next experiment we test this hypothesis.

In Table 7 we address the following question: If we have collected data only once, i.e. $n = 1$ in Theorem 3.1, which is often the case in applications, can the asymptotic result of Theorem 3.1 provide a reasonable estimate for the variability in the corresponding reconstruction $\hat{\epsilon}_{s,1}$ of ϵ_s^*? We perform the experiment for eight noise levels, and compute the corresponding maximum likelihood estimate $\hat{\epsilon}_{s,1}$. We approximate the standard deviation of $\hat{\epsilon}_{s,1}$ as in Table 6.

First, we note that as the noise level rises, so does the variability in the corresponding reconstruction $\hat{\epsilon}_{s,1}$. Furthermore, for each noise level, the reconstructions are well within two standard deviations of the true value ϵ_s^*. (See Table 4 for reconstructions with $n = 1$ for different data and at the same noise levels.) Most of the reconstructions, in fact, are near to or are within one standard deviation of ϵ_s^*. We conclude therefore that for this particular problem, the approximations of the standard deviation found in Table 7 provide one with a reasonable notion of the variability in the reconstructions $\hat{\epsilon}_{s,1}$.

Finally, from the above information we can obtain confidence intervals for our reconstruction $\hat{\epsilon}_{s,n}$. For this we use the Wald test (p. 525 of [9]), which says that under the regularity conditions of Theorem 3.1 we have that

$$(\hat{\epsilon}_{s,n} - \epsilon_s^*)\sqrt{n\hat{I}_n} \to \mathcal{N}(0, 1), \tag{62}$$

where $\hat{I}_n$ is a consistent estimator of $I(\epsilon_s^*)$. Then

$$\hat{\epsilon}_{s,n} - \frac{u_{\beta/2}}{\sqrt{n\hat{I}_n}} < \epsilon_s^* < \hat{\epsilon}_{s,n} + \frac{u_{\beta/2}}{\sqrt{n\hat{I}_n}}. \tag{63}$$

are confidence intervals for ϵ_s^* with asymptotic confidence coefficient $1 - \beta$, and $u_{\beta/2}$ is

Table 7. Estimates for the variability in $\hat{\epsilon}_s$ for $n = 1$. The values of σ, τ, and ϵ_∞ are fixed at the values found in Figure 4. We use α_n to estimate $I(\epsilon_s^*)$. Then $1/n\alpha_n$ estimates the variance and $1/\sqrt{n\alpha_n}$ estimates the standard deviation in (56).

% noise	$\hat{\epsilon}_{s,n}$	α_n	$1/n\alpha_n$	$1/\sqrt{n\alpha_n}$
0.5	80.37	6.6239	0.1510	0.3885
1.0	80.93	1.6242	0.6157	0.7847
1.5	81.50	0.7085	1.4115	1.1881
2.0	82.07	0.3906	2.5604	1.6001
2.5	82.65	0.2447	4.0867	2.0216
3.0	83.23	0.1664	6.0092	2.4514
5.0	85.63	0.0550	18.1739	4.2631

Table 8. Estimates for the Variability in $\hat{\epsilon}_s$ for the Example with 2% Noise. The values of σ, τ, and ϵ_∞ are fixed at the values found in Figure 4.

n	$\hat{\epsilon}_{s,n}$	$1/\sqrt{n\alpha_n}$	$\hat{\epsilon}_{s,n} - \frac{u_{\beta/2}}{\sqrt{n\alpha_n}}$	$\hat{\epsilon}_{s,n} + \frac{u_{\beta/2}}{\sqrt{n\alpha_n}}$
1	82.07	1.6001	77.95	86.19
2	80.36	1.0982	77.53	83.19
4	79.84	0.7705	77.86	81.82
8	80.38	0.5496	78.96	81.80
16	80.11	0.3868	79.11	81.11
32	80.01	0.2729	79.31	80.71
64	79.73	0.1921	79.23	80.22

the upper $\beta/2$ point of the standard normal distribution. It is given in [9], p. 526, that α_n in (60) is a consistent estimator of $I(\epsilon_s^*)$, and hence we can use $\hat{I}_n = \alpha_n$ in (63) to obtain confidence interval estimates.

For the results summarized in Table 6, we use (63) to obtain approximate confidence intervals for $\beta = .99$. The results are found in Table 8. We see that for n small, the confidence intervals are large, but as n becomes large, the confidence intervals become small, and a great deal more certainty can be inferred about the value of the reconstruction. We note that (63) is an asymptotic result, and hence, is suspect in the case of small n. Nonetheless, for each n in our example, the true value ϵ_s^* is contained within the given confidence interval.

For the results summarized in Table 7, we again use (63) to obtain approximate confidence intervals for $\beta = .9$. The results are found in Table 9. The true value ϵ_s^* is contained in the confidence interval for each noise level. Hence, we conclude that even for $n = 1$ the confidence intervals given by the Wald test (63) give a reasonable notion of the variability in the reconstructions of ϵ_s^*. Note the dramatic increase in the size of the confidence intervals as the noise level increases.

Table 9. Estimates for the variability in $\hat{\epsilon}_s$ for $n = 1$. The values of σ, τ, and ϵ_∞ are fixed at the values found in Figure 4. We use α_n to estimate $I(\epsilon_s^*)$. Then $1/n\alpha_n$ estimates the variance and $1/\sqrt{n\alpha_n}$ estimates the standard deviation in (56).

% noise	$\hat{\epsilon}_{s,n}$	$1/\sqrt{n\alpha_n}$	$\hat{\epsilon}_{s,n} - \frac{u_{\beta/2}}{\sqrt{n\alpha_n}}$	$\hat{\epsilon}_{s,n} + \frac{u_{\beta/2}}{\sqrt{n\alpha_n}}$
0.5	80.37	0.3885	79.73	81.01
1.0	80.93	0.7847	79.64	82.22
1.5	81.50	1.1881	79.55	83.45
2.0	82.07	1.6001	79.44	84.70
2.5	82.65	2.0216	79.32	85.98
3.0	83.23	2.4514	79.20	87.26
5.0	85.63	4.2631	78.62	92.64

5. Conclusions

We have presented a mathematical model for two dimensional, time domain, TM mode calculations with PML absorbing boundary conditions, in the presence of a Debye medium. The resulting set of PDEs is solved using the FDTD algorithm. We present numerical results from forward calculations when the Debye medium parameters are chosen in order to model the polarization behavior of water.

A statistical model for data generation is presented, and we solve the parameter identification problem by obtaining the maximum likelihood estimate corresponding to the data and noise model. For us this involves the minimization of the negative log-likelihood function, which takes a least squares form. A constrained, ellipsoidal, trust region modification of the Levenburg-Marquardt algorithm is used to this end.

A frequency domain analysis is presented that provides an explanation of results obtained from solutions of the inverse problem for data generated with varying levels of noise. Strategies for solving the inverse problem which decrease computational cost and are based on these conclusions are presented.

Finally, we use a theorem from the statistical theory of large samples to obtain estimates of the variability in the estimated parameters. Confidence intervals for these estimated parameters are also presented.

6. Acknowledgements

References

[1] Richard Albanese, John Penn, and Richard Medina. *Short-rise-time microwave pulse propagation through dispersive biological media.* J. Optical Soc. Amer. A, **6**(1989),pp. 1441-1446.

[2] H. T. Banks, Brian L. Browning. *Time domain electromagnetic scattering using perfectly matched layers.* CRSC Tech. Report TR02-23, NC State Univ., July 2002, Revised, June 2003; Comp. Meth. Appl. Mech. Engr., submitted.

[3] H. T. Banks, M. W. Buksas, and T. Lin. *Electromagnetic Material Interrogation Using Conductive Interfaces and Acoustic Wavefronts.* SIAM, Philadelphia, 2000.

[4] H. T. Banks and K. Kunisch. *Estimation Techniques for Distributed Parameter Systems.* Birkhauser, Boston, 1989.

[5] J. G. Blaschak and J. Franzen. *Precursor propagation in dispersive media from short-rise-time pulses at oblique incidence.* J. Optical Soc. Amer. A, **12** (1995), pp. 1501-1512.

[6] W. C. Chew and J. M. Jin. *Perfectly Matched Layers in the Discretized Space: An Analysis and Optimization.* Electromagnetics, **16** (1996), pp. 325-340.

[7] Stephen D. Gedney. *An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD lattices.* IEEE Trans. Antennas and Propagation, **44** (1996), pp. 1630-1639.

[8] Jianming Jin. *The Finite Element Method in Electromagnetics.* John Wiley and Sons, Inc., New York, 2002.

[9] E. L. Lehman *Elements of Large-Sample Theory.* Springer, New York, 1999.

[10] Hao Ling, Department of Electrical and Computer Engineering, University of Texas, private correspondence, 2/25/2003.

[11] Jorge Nocedal and Stephen Wright. *Numerical Optimization.* Springer, New York, 2000.

[12] Peter G. Petropoulos. *Stability and phase error analysis of FDTD in dispersive dielectrics.* IEEE Trans. Antennas and Propagation, **42** (1994), pp. 62-69.

[13] David M. Pozar. *Microwave Engineering.* John Wiley and Sons, Inc., New York, 1998.

[14] Z. S. Sacks, D. M. Kingsland, R. Lee, J. F. Lee. *A perfectly matched anisotropic absorber for use as an absorbing boundary condition.* IEEE Trans. Antennas and Propagation, **43**(1995), pp. 184-186.

[15] Dennis M. Sullivan. *Electromagnetic Simulation Using the FDTD Method.* IEEE Press, Piscataway, NJ, 2000.

[16] Allen Taflove and Susan C. Hagness. *Computational Electrodynamics: The Finite Difference Time Domain Method.* Artech House, Cambridge, 2000.